

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

**Ancient Ethiopian genome reveals extensive Eurasian admixture throughout the African continent**

**This is a pre print version of the following article:**

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/1661894> since 2018-03-11T22:28:49Z

*Published version:*

DOI:10.1126/science.aad2879

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

# **Title: Ancient Ethiopian genome reveals extensive Eurasian admixture throughout the African continent**

**Authors:** Gallego Llorente, M.<sup>1 \*†</sup>, Jones, E. R.<sup>2 \*†</sup>, Eriksson, A.<sup>1, 3</sup>, Siska, V.<sup>1</sup>, Arthur, K.W.<sup>4</sup>, Arthur, J.W.<sup>4</sup>, Curtis, M. C.<sup>5, 6</sup>, Stock, J.T.<sup>7</sup>, Coltorti, M.<sup>8</sup>, Pieruccini, P.<sup>8</sup>, Stretton, S.<sup>9</sup>, Brock, F.<sup>10, 11</sup>, Higham, T.<sup>10</sup>, Park, Y.<sup>12</sup>, Hofreiter, M.<sup>13, 14</sup>, Bradley, D. G.<sup>2</sup>, Bhak, J.<sup>15</sup>, Pinhasi, R.<sup>16 \*</sup>, Manica, A.<sup>1 \*</sup>

## **Affiliations:**

1. Department of Zoology, University of Cambridge, Downing Street, Cambridge, CB2 3EJ, UK.
2. Smurfit Institute of Genetics, Trinity College Dublin, Dublin, Ireland.
3. Integrative Systems Biology Laboratory, King Abdullah University of Science and Technology (KAUST), Thuwal, 23955-6900, Kingdom of Saudi Arabia.
4. Department of Society, Culture, and Language, University of South Florida St. Petersburg, 140 7th Avenue South, St. Petersburg, Florida 33701, USA.
5. Department of Anthropology, Ventura College, 4667 Telegraph Road, Ventura, California 93003, USA.
6. Humanities and Social Sciences Program, UCLA Extension, University of California Los Angeles, 10995 Le Conte Avenue, Los Angeles, California 90095, USA.
7. Department of Archaeology and Anthropology, University of Cambridge, Pembroke Street, Cambridge, CB2 3QG, UK.
8. Department of Physical Sciences, Earth and Environment, University of Siena, Via di Laterina, 8 - 53100 Siena, Italy.
9. Department of Anthropology, University of Illinois at Urbana-Champaign, Public Service Archaeology and Architecture Program, 109 Davenport Hall, 607 S. Mathews Ave. Urbana, Illinois 61801, USA.
10. Oxford Radiocarbon Accelerator Unit, Research Laboratory for Archaeology and the History of Art, University of Oxford, Dyson Perrins Building, South Parks Road, OX1 3QY, Oxford, UK.
11. Cranfield Forensic Institute, Cranfield University, Defence Academy of the United Kingdom, Shrivenham, Oxon, SN6 8LA, UK.
12. Theragen BiO Institute, 2nd Floor B-dong, AICT bldg, Iui-dong, Youngtong-gu, Suwon, 443-270, Republic of Korea.
13. Institute for Biochemistry and Biology, Faculty for Mathematics and Natural Sciences, University of Potsdam, Karl-Liebknechtstraße 24–25, 14476 Potsdam Golm, Germany.
14. Department of Biology, University of York, Wentworth Way, Heslington, York, YO10 5DD, UK.
15. The Genomics Institute, UNIST, Ulsan, 689-798, Republic of Korea.

16. School of Archaeology, University College Dublin, Belfield, Dublin 4, Ireland.

† These authors contributed equally to the work.

\* Correspondence to: [mg632@cam.ac.uk](mailto:mg632@cam.ac.uk) (M.G.); [joneser@tcd.ie](mailto:joneser@tcd.ie) (E.R.J.); [ron.pinhasi@ucd.ie](mailto:ron.pinhasi@ucd.ie) (R.P.); [am315@cam.ac.uk](mailto:am315@cam.ac.uk) (A.M.)

**Abstract:** Characterizing genetic diversity in Africa is a crucial step for most analyses reconstructing the evolutionary history of anatomically modern humans. However, historic migrations from Eurasia into Africa have affected many contemporary populations, confounding inferences. Here, we present a 12.5x coverage ancient genome of an Ethiopian male ('Mota') who lived approximately 4,500 years ago. We use this genome to demonstrate that the Eurasian backflow into Africa came from a population closely related to Early Neolithic farmers, who had colonized Europe 4,000 years earlier. The extent of this backflow was much greater than previously reported, reaching all the way to Central, West and Southern Africa, affecting even populations such as Yoruba and Mbuti, previously thought to be relatively unadmixed, who harbor 6-7% Eurasian ancestry.

**One Sentence Summary:** An ancient genome from Ethiopia reveals widespread Eurasian admixture into modern African populations.

**Main Text:** The ability to sequence ancient genomes has revolutionized our understanding of human evolution. However, genetic analyses of ancient material have focused on individuals from temperate and arctic regions, where ancient DNA is preserved over longer time frames (1). Africa has so far failed to yield skeletal remains with much aDNA, with the exception of a few poorly preserved specimens from which only mitochondrial DNA could be extracted (2). This is particularly unfortunate, as African genetic diversity is crucial to most analyses reconstructing the evolutionary history of anatomically modern humans, by providing the baseline against which other events are defined. In the absence of ancient DNA, geneticists rely on contemporary African populations, but a number of historic events, in particular a genetic backflow from West Eurasia into Eastern Africa (3, 4), act as confounding factors.

Here, we present an ancient human genome from Africa, and use it to disentangle the effects of recent population movement into Africa. By sampling the petrous bone (5), we sequenced the genome of a male from Mota Cave (herein referred to as 'Mota') in the southern Ethiopian highlands, with a mean coverage of 12.5x (6). Contamination was estimated to be between 0.29 and 1.26% (6). Mota's remains were dated to ~4,500 years ago (direct calibrated radiocarbon date (6)), and thus predate both the Bantu expansion (7), and, more importantly, the 3ky-old West Eurasian backflow which has left strong genetic signatures in the whole of Eastern and, to a lesser extent, Southern Africa (3, 4).

We compared Mota to contemporary human populations (6). Both Principal Component Analysis (Fig. 1A) and outgroup  $f_3$  analysis using Ju|'hoansi (Khoisan) from Southern Africa as the outgroup (Fig. 1B,C) place this ancient individual close to contemporary Ethiopian populations, and more specifically to the Ari, a group of Omotic speakers from southern Ethiopia, to the West of the highland region where Mota lived. Our ancient genome confirms the view that the divergence of this language family results from the relative isolation of its speakers (8), and indicates population continuity over the last ~4,500 years in this region of Eastern Africa.

The age of Mota means that he should predate the West Eurasian backflow, which has been dated to ~3,000 years ago (3, 4). We formally tested this by using an  $f_4$  ratio estimating the West Eurasian component (6), following the approach adopted by Pickrell et al (3). As expected, we failed to find any West Eurasian component in Mota (Table S5), thus providing support for previous dating of that event (3, 4).

Given that Mota predates the backflow, we searched for its most likely source by modelling the Ari, the contemporary population closest to our ancient genome, as a mixture of Mota and another West Eurasian population (6). We investigated both contemporary sources (3) as well as other Eurasian ancient genomes (5, 9). In this analysis, contemporary Sardinians and the early Neolithic LBK (Stuttgart) genome stand out (Fig. 2A). Previous analyses have shown Sardinians to be the closest modern representatives of early Neolithic farmers (10, 11), implying that the backflow came from the same genetic source that fuelled the Neolithic expansion into Europe from the Near East/Anatolia, before recent historic events changed the genetic makeup of populations living in that region. An analysis with haplotype sharing also identified a connection between contemporary Ethiopians and Anatolia (4, 12). Interestingly, archaeological evidence dates the arrival of Near Eastern domesticates (such as wheat, barley and lentils) to the same time period (circa 3,000 years ago) (13, 14), suggesting that the direct descendants of the farmers that earlier brought agriculture into Europe may have also played a role in the development of new forms of food production in the Horn of Africa.

Using Mota as an unadmixed African reference and the early farmer LBK as the source of the West Eurasian component, it is possible to reassess the magnitude and geographic extent of historical migrations, avoiding the complications of using admixed contemporary populations (6). We estimated a substantially higher Eurasian backflow admixture than previously detected (3), with an additional 4-7% of the genome of most African populations tracing back to a Eurasian source, and, more importantly, we detected a much broader geographical impact of the backflow, going all the way to West and Southern Africa (Fig. 2B). Even though the West Eurasian component in these regions is smaller than in Eastern Africa, it is still sizeable, with Yoruba and Mbuti, who are often used

as African references (15, 16), showing 7% and 6%, respectively, of their genomes to be of Eurasian origin (Table S5).

Since Mota predates recent demographic events, his genome can act as an ideal African reference to understand episodes during the out-of-Africa expansion. We used him as the African reference to quantify Neanderthal introgression in a number of contemporary genomes (6). Both Yoruba and Mbuti, which are routinely used as African references for this type of analysis (15, 16), show a marginally closer affinity with Neanderthal than Mota based on  $D$  statistics, and an  $f_4$  ratio analysis detected a small Neanderthal component in these genomes at around 0.2-0.7%; greater than previously suggested (16), and consistent with our estimates of the magnitude of their Western Eurasian ancestry (6). Whilst the magnitude of Neanderthal ancestry in these contemporary African populations is not enough to change conclusions qualitatively (estimates of Neanderthal ancestry in French and Han only increased marginally when tested with Mota as a reference), it should be accounted for when looking for specific introgressed haplotypes (17) or searching for unknown ancient hominins who might have hybridized with African populations (18).

We also investigated the Mota genome for a number of phenotypes of interest (6). As expected, Mota lacked any of the derived alleles found in Eurasian populations for eye and skin colour, suggesting that he had brown eyes and dark skin. Mota lacked any of the currently known alleles that give lactose tolerance, which may have implications concerning when pastoralism appeared in southwestern Ethiopia. In addition, Mota did possess all three selected alleles that have been recently shown to play a role in the adaptation to altitude in contemporary highland Ethiopian populations (19). The presence of these mutations supports our conclusion that Mota is the descendant of highland dwellers, who have lived in this environment long enough to accumulate adaptations to the altitude (20, 21).

Until now, it has been necessary to use contemporary African populations as the baseline against which events during the worldwide expansion of Anatomically Modern Humans are defined (16, 22–24). By obtaining an ancient whole genome from this continent, we have shown that having an unadmixed reference that predates the large number of recent historical migrations can greatly improve our inference. This result stresses the importance of obtaining unadmixed baseline data to reconstruct demographic events, and the limitations of analyses that are solely based on contemporary populations. Even older African genomes will thus be needed to investigate key demographic events that predate Mota, such as earlier instances of back-flows into Africa (25).

## References and Notes:

1. M. Hofreiter *et al.*, The future of ancient DNA: Technical advances and conceptual shifts. *Bioessays*. **37**, 284–293 (2015).
2. A. G. Morris, A. Heinze, E. K. F. Chan, A. B. Smith, V. M. Hayes, First Ancient Mitochondrial Human Genome from a Prepastoralist Southern African. *Genome Biol. Evol.* **6**, 2647–2653 (2014).
3. J. K. Pickrell *et al.*, Ancient west Eurasian ancestry in southern and eastern Africa. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2632–2637 (2014).
4. L. Pagani *et al.*, Ethiopian genetic diversity reveals linguistic stratification and complex influences on the Ethiopian gene pool. *Am. J. Hum. Genet.* **91**, 83–96 (2012).
5. C. Gamba *et al.*, Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* **5** (2014), doi:10.1038/ncomms6257.
6. See supporting material on Science Online.
7. S. Li, C. Schlebusch, M. Jakobsson, Genetic variation reveals large-scale population expansion and migration during the expansion of Bantu-speaking peoples. *Proc. Biol. Sci.* **281** (2014), doi:10.1098/rspb.2014.1448.
8. R. Blench, in *SemitoHamitic Festschrift for A.B. Dolgopolsky and H. Jungrauthmayr* (Dietrich Reimer Verlag, Berlin, 2008), pp. 63–78.
9. I. Lazaridis *et al.*, Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*. **513**, 409–413 (2014).
10. M. Sikora *et al.*, Population Genomic Analysis of Ancient and Modern Genomes Yields New Insights into the Genetic Ancestry of the Tyrolean Iceman and the Genetic Structure of Europe. *PLoS Genet.* **10**, e1004353 (2014).
11. P. Skoglund *et al.*, Origins and Genetic Legacy of Neolithic Farmers and Hunter-Gatherers in Europe. *Science*. **336**, 466–469 (2012).
12. T. Kivisild *et al.*, Ethiopian Mitochondrial DNA Heritage: Tracking Gene Flow Across and Around the Gate of Tears. *Am. J. Hum. Genet.* **75**, 752–770 (2004).
13. Curtis, M.C., in *The Oxford Handbook of African Archaeology* (Oxford University Press, 2013), pp. 571–584.
14. M. Harrower, M. McCriston, A. D’Andrea, General/Specific, Local/Global: Comparing the Beginnings of Agriculture in the Horn of Africa (Ethiopia/Eritrea) and Southwest Arabia (Yemen). *Am. Antiq.* **75**, 452–472 (2010).
15. Q. Fu *et al.*, An early modern human from Romania with a recent Neanderthal ancestor. *Nature*. **524**, 216–219 (2015).

16. K. Prüfer *et al.*, The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*. **505**, 43–49 (2014).
17. S. Sankararaman *et al.*, The genomic landscape of Neanderthal ancestry in present-day humans. *Nature*. **507**, 354–357 (2014).
18. M. F. Hammer, A. E. Woerner, F. L. Mendez, J. C. Watkins, J. D. Wall, Genetic evidence for archaic admixture in Africa. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15123–15128 (2011).
19. N. Udpa *et al.*, Whole genome sequencing of Ethiopian highlanders reveals conserved hypoxia tolerance genes. *Genome Biol.* **15**, R36 (2014).
20. E. Huerta-Sánchez *et al.*, Genetic Signatures Reveal High-Altitude Adaptation in a Set of Ethiopian Populations. *Mol. Biol. Evol.* **30**, 1877–1888 (2013).
21. G. Alkorta-Aranburu *et al.*, The Genetic Architecture of Adaptations to High Altitude in Ethiopia. *PLoS Genet.* **8**, e1003110 (2012).
22. A. Eriksson *et al.*, Late Pleistocene climate change and the global expansion of anatomically modern humans. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 16089–16094 (2012).
23. A. Eriksson, A. Manica, Effect of ancient population structure on the degree of polymorphism shared between modern human populations and ancient hominins. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 13956–13960 (2012).
24. L. Pagani *et al.*, Tracing the Route of Modern Humans out of Africa by Using 225 Human Genome Sequences from Ethiopians and Egyptians. *Am. J. Hum. Genet.* **96**, 986–991 (2015).
25. J. A. Hodgson, C. J. Mulligan, A. Al-Meerri, R. L. Raaum, Early Back-to-Africa Migration into the Horn of Africa. *PLoS Genet.* **10**, e1004393 (2014).
26. F. Brock, T. Higham, P. Ditchfield, C. B. Ramsey, Current Pretreatment Methods for AMS Radiocarbon Dating at the Oxford Radiocarbon Accelerator Unit (ORAU). *Radiocarbon*. **52**, 103–112 (2010).
27. J. E. Buikstra, D. H. Ubelaker, Standards for data collection from human skeletal remains (1994), vol. 44 of Arkansas Archaeological Survey Research Series.
28. M. R. Fedelsman, R. L. Fountain, “Race” Specificity and the Femur/Stature Ratio. *Am. J. Phys. Anthropol.* **100**, 207–224 (1996).
29. M. H. Raxter *et al.*, Stature estimation in ancient Egyptians: A new technique based on anatomical reconstruction of stature. *Am. J. Phys. Anthropol.* **136**, 147–155 (2008).
30. C. B. Ruff, E. Trinkaus, T. W. Holliday, Body mass and encephalization in Pleistocene Homo. *Nature*. **387**, 173–176 (1997).
31. R. Pinhasi *et al.*, Optimal Ancient DNA Yields from the Inner Ear Part of the Human Petrous Bone. *PLOS ONE*. **10**, e0129102 (2015).



32. D. Y. Yang, B. Eng, J. S. Waye, J. C. Dudar, S. R. Saunders, Technical note: improved DNA extraction from ancient bones using silica-based spin columns. *Am. J. Phys. Anthropol.* **105**, 539–543 (1998).
33. M. Meyer, M. Kircher, Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. *Cold Spring Harb. Protoc.* **2010**, pdb.prot5448 (2010).
34. G. Renaud, U. Stenzel, J. Kelso, leeHom: adaptor trimming and merging for Illumina sequencing reads. *Nucleic Acids Res.* **42**, e141 (2014).
35. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* **25**, 1754–1760 (2009).
36. A. McKenna *et al.*, The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
37. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* **26**, 841–842 (2010).
38. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* **25**, 2078–2079 (2009).
39. H. Jónsson, A. Ginolhac, M. Schubert, P. L. F. Johnson, L. Orlando, mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics.* **29**, 1682–1684 (2013).
40. A. W. Briggs *et al.*, Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 14616–14621 (2007).
41. P. Brotherton *et al.*, Novel high-resolution characterization of ancient DNA reveals C > U-type base modification events as the sole cause of post mortem miscoding lesions. *Nucleic Acids Res.* **35**, 5717–5728 (2007).
42. B. Shapiro, M. Hofreiter, A Paleogenomic Perspective on Evolution and Gene Function: New Insights from Ancient DNA. *Science.* **343**, 1236573 (2014).
43. T. S. Korneliussen, A. Albrechtsen, R. Nielsen, ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinform.* **15**, 356 (2014).
44. M. Rasmussen *et al.*, An Aboriginal Australian Genome Reveals Separate Human Dispersals into Asia. *Science.* **334**, 94–98 (2011).
45. F. Sánchez-Quinto *et al.*, Genomic Affinities of Two 7,000-Year-Old Iberian Hunter-Gatherers. *Curr. Biol.* **22**, 1494–1499 (2012).
46. D. Vianello *et al.*, HAPLOFIND: a new method for high-throughput mtDNA haplogroup assignment. *Hum. Mutat.* **34**, 1189–1194 (2013).
47. P. Skoglund, J. Storå, A. Götherström, M. Jakobsson, Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J. Archaeol. Sci.* **40**, 4477–4482 (2013).

48. P. Skoglund *et al.*, Genomic Diversity and Admixture Differs for Stone-Age Scandinavian Foragers and Farmers. *Science*. **344**, 747–750 (2014).
49. D. M. Behar *et al.*, A “Copernican” reassessment of the human mitochondrial DNA tree from its root. *Am. J. Hum. Genet.* **90**, 675–684 (2012).
50. P. Soares *et al.*, The expansion of mtDNA haplogroup L3 within and out of Africa. *Mol. Biol. Evol.* **29**, 915–927 (2012).
51. A. Torroni, A. Achilli, V. Macaulay, M. Richards, H.-J. Bandelt, Harvesting the fruit of the human mtDNA tree. *Trends Genet.* **22**, 339–345 (2006).
52. D. M. Behar *et al.*, The dawn of human matrilineal diversity. *Am. J. Hum. Genet.* **82**, 1130–1140 (2008).
53. L. Jostins *et al.*, YFitter: Maximum likelihood assignment of Y chromosome haplogroups from low-coverage sequence data. *arXiv:1407.7988 [q-bio]* (2014) (available at <http://arxiv.org/abs/1407.7988>).
54. O. Semino *et al.*, Origin, Diffusion, and Differentiation of Y-Chromosome Haplogroups E and J: Inferences on the Neolithization of Europe and Later Migratory Events in the Mediterranean Area. *Am. J. Hum. Genet.* **74**, 1023–1034 (2004).
55. B. Trombetta, F. Cruciani, D. Sellitto, R. Scozzari, A New Topology of the Human Y Chromosome Haplogroup E1b1 (E-P2) Revealed through the Use of Newly Characterized Binary Polymorphisms. *PLOS ONE*. **6**, e16073 (2011).
56. E. I. Gebremeskel, M. E. Ibrahim, Y-chromosome E haplogroups: their distribution and implication to the origin of Afro-Asiatic languages and pastoralism. *Eur. J. Hum. Genet.* **22**, 1387–1392 (2014).
57. O. Semino, A. S. Santachiara-Benerecetti, F. Falaschi, L. L. Cavalli-Sforza, P. A. Underhill, Ethiopians and Khoisan share the deepest clades of the human Y-chromosome phylogeny. *Am. J. Hum. Genet.* **70**, 265–268 (2002).
58. J. K. Pickrell *et al.*, The genetic prehistory of southern Africa. *Nat. Commun.* **3**, 1143 (2012).
59. P. Danecek *et al.*, The variant call format and VCFtools. *Bioinformatics*. **27**, 2156–2158 (2011).
60. S. Purcell *et al.*, PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
61. N. Patterson, A. L. Price, D. Reich, Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).
62. L. van Dorp *et al.*, Evidence for a Common Origin of Blacksmiths and Cultivators in the Ethiopian Ari within the Last 4500 Years: Lessons for Clustering-Based Inference. *PLoS Genet.* **11**, e1005397 (2015).
63. N. J. Patterson *et al.*, Ancient Admixture in Human History. *Genetics*. **192**, 1065–1093 (2012).

64. P. Moorjani *et al.*, The History of African Gene Flow into Southern Europeans, Levantines, and Jews. *PLoS Genet.* **7**, e1001373 (2011).
65. V. Fernandes *et al.*, Genetic stratigraphy of key demographic events in Arabia. *PLOS ONE.* **10**, e0118625 (2015).
66. R. E. Green *et al.*, A Draft Sequence of the Neandertal Genome. *Science.* **328**, 710–722 (2010).
67. M. Meyer *et al.*, A High-Coverage Genome Sequence from an Archaic Denisovan Individual. *Science.* **338**, 222–226 (2012).
68. I. V. Ovchinnikov *et al.*, Molecular analysis of Neanderthal DNA from the northern Caucasus. *Nature.* **404**, 490–493 (2000).
69. K. L. Hart *et al.*, Improved eye- and skin-color prediction based on 8 SNPs. *Croat. Med. J.* **54**, 248–256 (2013).
70. S. Walsh *et al.*, The HirisPlex system for simultaneous prediction of hair and eye colour from DNA. *Forensic Sci. Int. Genet.* **7**, 98–115 (2013).
71. T. G. K. Jensen *et al.*, The -14010\*C variant associated with lactase persistence is located between an Oct-1 and HNF1 $\alpha$  binding site and increases lactase promoter activity. *Hum. Genet.* **130**, 483–493 (2011).
72. S. A. Tishkoff *et al.*, Convergent adaptation of human lactase persistence in Africa and Europe. *Nature Genet.* **39**, 31–40 (2007).
73. B. L. Jones *et al.*, Diversity of Lactase Persistence Alleles in Ethiopia: Signature of a Soft Selective Sweep. *Am. J. Hum. Genet.* **93**, 538–544 (2013).
74. S. Bergink, S. Jentsch, Principles of ubiquitin and SUMO modifications in DNA repair. *Nature.* **458**, 461–467 (2009).

**Acknowledgements:** A.M. was supported by ERC Consolidator Grant 647787 ‘LocalAdaptation’; R.P. by ERC Starting Grant: ERC- 2010-StG 26344; M.H. by ERC Consolidator Grant 310763 ‘GeneFlow’; J.B. by the 2014 Research Fund (1.140113.01, 1.140064.01) of UNIST (Ulsan National Institute of Science and Technology) and Geromics internal research funding; J.T.S. by ERC Consolidator Grant 617627 ‘ADaPt’; K.W.A. by NSF award 1027607; D.G.B. by ERC Investigator Grant 295729-CodeX; V.S. by a scholarship from the Gates Cambridge Trust; and M.G. by a BBSRC DTP studentship. Permission for the archaeology was given by the Ethiopian Authority for Research and Conservation of Cultural Heritage and offices of the Ministry of Culture and Tourism for the Southern Nations, Nationalities, and Peoples Region. Raw reads from Mota are available for download through NCBI, BioProject ID PRJNA295861, and the corresponding BAM and VCF files are available at [africangenome.org](http://africangenome.org).

## Figure legends

**Fig. 1. Mota shows a very high degree of similarity with the highland Ethiopian Ari populations.** (A), PCA showing Mota projected onto components loaded on contemporary African and Eurasian populations. The inset magnifies the PCA space occupied by Ethiopian and Eastern African populations. (B), outgroup  $f_3$  quantifying the shared drift between Mota and contemporary African populations, using Ju|'hoansi (Khoisan) as an outgroup; bars represent standard error; (C), map showing the distribution of outgroup  $f_3$  values across the African continent. In (A) and (B), populations speaking Nilo-Saharan languages are marked with blue shades, Omotic speakers with red, Cushitic with orange, Semitic with yellow, and Bantu with green. Mota is denoted by a black symbol.

**Fig. 2. Quantifying the geographic extent and origin of the West Eurasian component in Africa.** (A), admixture  $f_3$  identifying likely sources of the West Eurasian component (lowest  $f_3$  values). Contemporary populations in blue, ancient genomes in red; bars represent standard error. (B), map showing the proportion of West Eurasian component,  $\lambda_{\text{Mota,LBK}}$ , across the African continent.

**Supplementary Materials:**

SOM Text

References 26-74

Figures S1-S8

Tables S1-S14